Formulation and Optimization of Robust Sensor Placement Problems for Drinking Water Contamination Warning Systems

Jean-Paul Watson[*] Regan Murray[†] William E. Hart[‡]

Abstract

The sensor placement problem in contamination warning system design for water distribution networks involves maximizing the protection level afforded by limited numbers of sensors, typically quantified as the expected impact of a contamination event; the issue of how to mitigate against high-impact events is either handled implicitly or ignored entirely. Consequently, such sensor placements run the risk of failing to protect against high-impact, 9/11-style attacks. In contrast, robust sensor placements address this concern by focusing strictly on high-impact events, and placing sensors to minimize the impact of these events. We introduce several robust formulations of the sensor placement problem, distinguished by how they quantify the potential damage due to high-impact events. Via solution of these formulations, we explore the nature of robust versus expectation-based sensor placements on three real-world, large-scale networks. We find that robust sensor placements can yield large reductions in the number and magnitude of high-impact events, for modest increases in expected impact. The resulting ability to trade-off between robust and expected-case impacts is a key, unexplored dimension in contamination warning system design.

Keywords

Sensor Placement, Contamination Warning System Design, Robust Optimization, Drinking Water, Homeland Security.

[*]Discrete Algorithms and Math Department, Sandia National Laboratories, P.O. Box 5800, MS 1318, Albuquerque, NM 87185-1318; E-Mail: `jwatson@sandia.gov`

[†]Environmental Protection Agency, NHSRC, Cincinnati, Ohio; E-Mail: `Murray.Regan@epamail.epa.gov`

[‡]Discrete Algorithms and Math Department, Sandia National Laboratories, P.O. Box 5800, MS 1318, Albuquerque, NM, 87185-1318; E-Mail: `wehart@sandia.gov`

1 Introduction

Contamination warning systems (CWSs) have been proposed as a promising approach for detecting contamination events in drinking water distribution systems. The goal of a CWS is to detect contamination events early enough to allow for effective public health and/or water utility intervention to limit potential public health or economic impacts. There are many challenges to detecting contaminants in drinking water systems: municipal distribution systems are large, consisting of hundreds or thousands of miles of pipe; flow patterns are driven by time-varying demands placed on the system by customers; and distribution systems are looped, resulting in mixing and dilution of contaminants. The drinking water community has proposed that CWSs be designed to maximize the number of contaminants that can be detected in drinking water distribution systems by combining online sensors with public health surveillance systems, physical security monitoring, customer complaint surveillance, and routine sampling programs (USEPA, 2005).

Algorithms for placing sensors to support the design of CWSs for municipal water distribution networks have received significant attention from researchers and practitioners over the last ten years (Kessler et al., 1998; Ostfeld and Salomons, 2004; Berry et al., 2005a, 2006b). Without exception, these algorithms attempt to either minimize the expected impact of a contamination event (e.g., in terms of the number of people sickened or the volume of contaminated water consumed) or maximize the proportion of contamination events that are ultimately detected, independent of impact. Recently, Berry et al. (2006b) showed that both objectives can be formulated in terms of a single optimization model, illustrating that the proportion of events detected can be viewed as an expected impact, and vice versa. In this unified optimization model, contamination event probabilities are either assumed to be uniform, or are estimated based on factors such as the difficulty of accessing a particular component of a distribution network. Given a broad range of possible contamination events, sensor placement algorithms then attempt to minimize the probability-weighted sum of contamination event impact, i.e., the expected impact. The most advanced algorithms currently available can successfully generate provably optimal sensor placements to very large (e.g., 10,000+ junction) distribution networks for very large numbers (e.g., 50,000+) of possible contamination events, in modest run-times on a modern computing workstation (Berry et al., 2006b). Consequently, the basic sensor placement problem for CWS design is largely solved for most practical distribution networks, and the research emphasis has moved toward the integration of more realistic modeling assumptions such as sensor failures (Berry et al., 2006a), site specific installation costs and accessibility considerations (Berry et al., 2005b), significantly larger numbers of possible contamination scenarios (Berry et al., 2007), and solution robustness in the face of data uncertainties (Carr et al., 2006).

One currently unexplored, but – we argue – critical aspect of the sensor placement problem involves optimization models in which the design objective is not minimization of the expected impact, but rather minimization of the worst-case impact or other "robust" measures that focus strictly on high-consequence contamination events. The lack of research into these alternative models is perhaps counterintuitive in a post-9/11 environment. One explanation is that most environmental problems have required a focus on mitigating all risks to human health, and not just associated with those extremely high-impact events. Yet, robust sensor placement optimization is of interest in practice. In our working experience with various US water municipalities, a common reaction when discussing expectation-based optimization models is "Why not *only* concentrate on high-impact contamination events?" Additional motivation for pursuing robust sensor placement models stems from the observation that even optimal expectation-based sensor placements can permit numerous high-impact contamination events (e.g., as discussed below in Section 2). Further, accurate estimation of event probabilities is notoriously difficult, allowing for unintended de-emphasis of high-impact events. Although the final determination of the design objective ultimately rests with policy-makers at various levels in disparate organizations, the aforementioned factors strongly suggest that, at a minimum, there is a need to understand the differences between expectation-based and robust sensor placements.

For example, one might conjecture that minimization of the worst-case impact to "acceptable" levels may require fewer overall sensors than minimization of the expected-case impact. However, our analyses support the opposite conclusion. Consider the illustrative scenario in which there exist n contamination events yielding impacts greater than some acceptable threshold T. Further assume that the n events target disparate regions of the network, such that a sensor will mitigate against only one of the n events. In such a situation, n sensors are required to achieve a worst-case impact below T. In contrast, only a small number of sensors $s < n$ may be necessary to yield significant reductions in mean impact, as those sensors are free to be placed at locations in the network capable of detecting contmainant from a broad range of events.

In this paper, we introduce a number of robust measures of sensor placement performance, drawing heavily from existing literature on robust optimization from the financial community. Using a variety of optimization algorithms, we construct sensor placements that minimize these robust impact measures on three real-world water distribution networks. We find that sensor placements that minimize the expected impact admit – without exception – a non-trivial number of very high-impact contamination events. These high-impact events can be mitigated with robust sensor placements, e.g., we observe that significant reductions in the worst-case impact are possible. These reductions come at the necessary expense of an increase in the mean impact of a contamination event. However, the degree to which trade-offs are possible is significantly larger than anticipated; performance discrepancies are so large that it is likely to influence the higher-level CWS design process. However, we show that the different robust measures are in fact largely uncorrelated, such that minimization of one measure can yield highly sub-optimal performance in terms of the other measures. Consequently, it is imperative for decision-makers to understand the nature of the various robust measures in detail.

The remainder of this paper is organized as follows. We begin in Section 2 with a motivating example to illustrate differences in the characteristics of sensor placements that are optimal with respect to expectation-based and worst-case performance. Various robust impact measures are then introduced in Section 3. Section 4 details the test networks, contamination event scenarios, optimization models, and algorithms that we use in the analysis discussed in Section 5; the latter details quantitative and qualitative differences between expectation-based and robust sensor placements. We defer discussion of the specific computational characteristics of the algorithms used in our analysis to Section 6, which additionally addresses the computational difficulty of robust sensor placement formulations. Finally, we conclude in Section 7 with a discussion of the implications of our results.

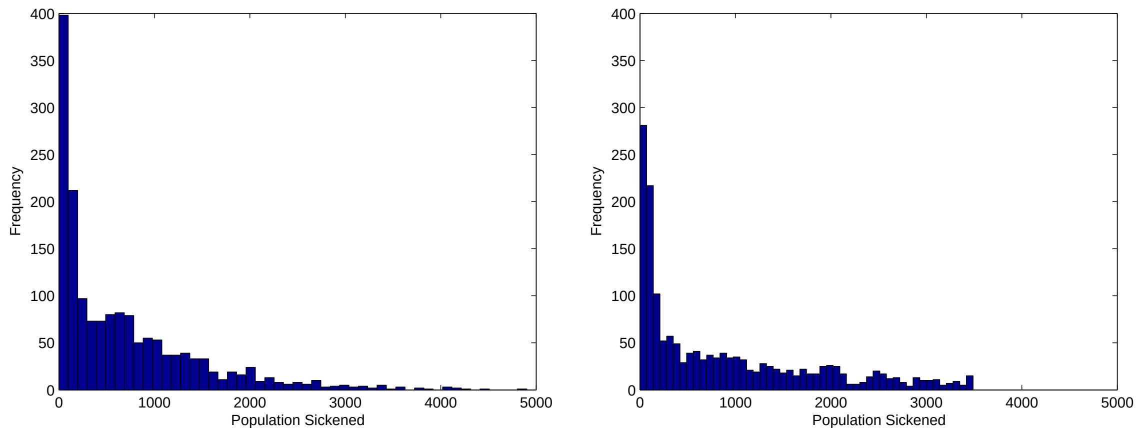

Figure 1: Histograms of the number of individuals sickened for various contamination events in Network2 under expected-case (left figure) and worst-case (right figure) sensor placements.

2 Motivating Example

To concretely illustrate the relative trade-offs that are possible between expectation-based and robust sensor placements, we begin with an example from a real-world water distribution network. The network is simply denoted Network2; this and other test networks are described in Section 4. Using the experimental methodology and algorithms presented below, we determine distinct sensor placements for Network2 – given a budget of 20 sensors – that respectively minimize the expected-case and worst-case impact of a contamination event. The precise details of the contamination scenarios are documented in Section 4; impact is quantified as the number of people sickened by a contamination event (Murray et al., 2006).

Histograms of the impacts of various contamination events (in this case, contaminant injections at each network node, for a total of approximately 1,600 events) *given* the expected-case and worst-case sensor placements are shown in Figure 1. We first consider the distribution of impacts under the expected-case sensor placement, as shown in the left side of Figure 1. The mean and worst-case impacts given this sensor placement are 685 and 4,902 individuals, respectively. The distribution exhibits a feature of sensor placements that minimize the expected-case: the presence of non-trivial numbers of contamination events that yield impacts over seven times greater than that of the mean. Specifically, eight contamination events yield impacts greater than 4,000 individuals sickened, while an additional six contamination events yields impacts between 3,500 and 4,000 individuals sickened.

Next, we consider the distribution of impacts given a sensor placement that minimizes the worst-case impact of a contamination event, as shown in the right side of Figure 1. In contrasting the two distributions, we immediately observe a significant reduction in the number of very high-impact contamination events. In particular, the highest-impact event sickens 3,490 individuals, in contrast to 4,902 individuals under the expected-case sensor placement; the 14 highest-impact events in the expected-case placement are mitigated by a sensor placement that minimizes the worst case. However, as is expected, the mitigation of high-impact events increases the frequency of small-to-moderate impact events. The worst-case sensor placement yields a mean impact of 882 individuals sickened, representing a 29% increase relative to the expected-case sensor placement. Even more dramatic growth is observed in the upper bound of the third impact quartile, from 1,011 under the expected-case sensor placement to 1,445 under the worst-case sensor placement (representing a 43% increase). The question for decision-makers in CWS design is then: Is a large (in this case 29%) reduction in the worst-case impact worth a correspondingly large increase in the expected impact?

A combination of adversarial and engineering factors ultimately dictate the answer to this question. Although very high-impact contamination events typically represent a small fraction of the total number of possible contamination events, they are not *a priori* any more difficult to initiate. Backflow injections can be carried out with roughly equiprobable success at most nodes in a typical distribution network. Further, contamination event probabilities are notoriously difficult to accurately quantify, due to a variety of estimates that must be made with respect to adversarial intent and capability, and the level of target vulnerability. Reliance on estimated event probabilities is therefore not without potentially significant risk; de-emphasis of high-impact contaminations with perceived low probability of occurrence may cause sensors to be placed in regions of the network that may allow many worst- or near-worst-case events to proceed unmitigated. Finally, adversarial characteristics have a significant impact on the design and assessment of a sensor placement. Intelligent and informed adversaries are likely to identify and initiate those contamination events that yield the highest-consequence impacts. Although some measures can be taken to mitigate intelligent adversaries (e.g., security classification of water system infrastructure and operations so that impacts cannot be easily predicted), they do not guarantee protection; insiders will always remain a threat, and trained engineers may be able to infer such characteristics from external observation or on-line resources with sufficient accuracy. Consequently, a rational and practical alternative to estimating contamination event probabilities is to simply assume an omniscient adversary and focus on protecting against worst-case events. This is a commonly adopted analysis approach when faced with adversarial conditions, and is routinely practiced in Operations Research and related decision support literature.

3 Quantifying Solution Robustness

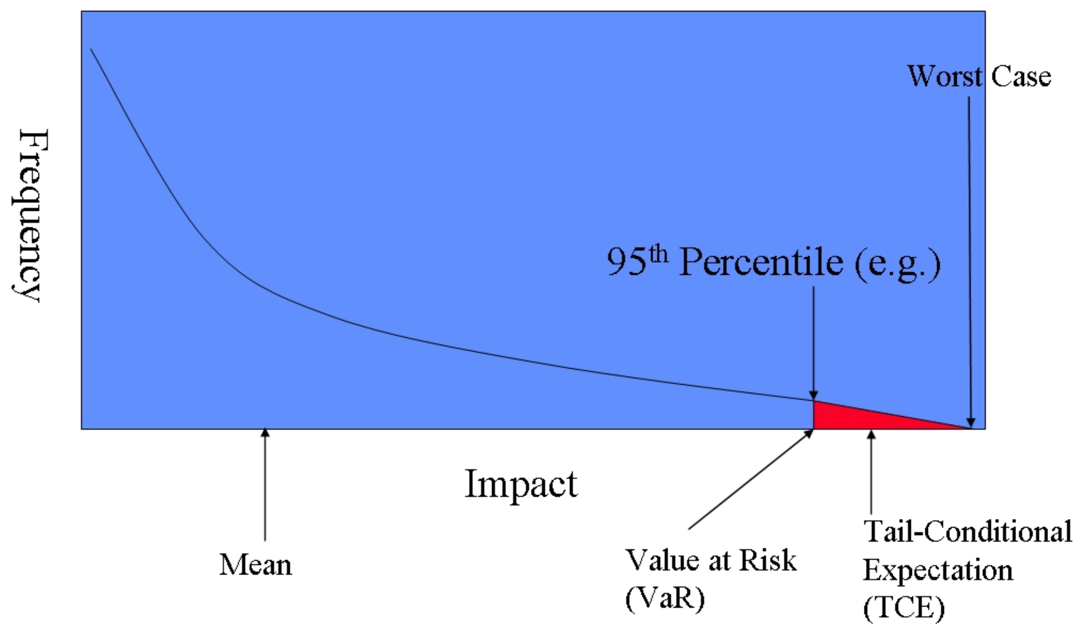

Figure 2: Graphical illustration contrasting the various "robust" metrics of sensor placement performance.

Informally, "robust" optimization methods focus on generating solutions that minimize down-side risk, i.e., the probability of occurrence of high-consequence events. The majority of early research on robust optimization originated in the academic financial community. Especially in this context, quantification of solution robustness is a key component of a robust optimization method. Two primary measures of solution robustness can be found in the body of financial literature: Value-at-Risk (*VaR*) and Tail-Conditional Expectation (*TCE*). Given a set of potential scenarios and their associated costs (e.g., impact to the population in the context of sensor placement), *VaR* is defined as the cost of the $(1-\alpha)\%$ most costly scenario (Holton, 2003), where $0 \leq \alpha \leq 1$. Typically, α is taken to be 0.05, such that the minimization of *VaR* effectively allows an optimization algorithm to ignore any costs associated with the $100 \cdot \alpha$ % highest-impact scenarios. *VaR* is an international standard for risk quantification in the banking community, and has seen widespread application in related contexts. In contrast to *VaR*, *TCE* quantifies the expected cost of the α most costly scenarios (Artzner et al., 1999); again, α is typically taken to be 0.05. Consequently, algorithms that minimize *TCE* must make decisions in order to reduce the tail mass of the cost distribution. The conditional value-at-risk measure, denoted *CVaR*, is closely related to the concept of *TCE*. In the case of continuous cost distributions, *CVaR* $=$ *TCE*. In the case of discrete cost distributions, *CVaR* is a continuous approximation to the true cost distribution, such that *TCE* $\leq$ *CVaR*. Finally, we additionally consider perhaps the most intuitive measure of down-side risk, that of the worst case cost, which we denote simply as *Worst*. Overall, we observe that these four risk or robustness measures are related through the following inequality: *VaR* $\leq$ *TCE* $\leq$ *CVaR* $\leq$ *Worst*. The various robust metrics are illustrated graphically in Figure 2.

4 Test Networks and Problem Formulation

We now describe the test networks (Section 4.1), experimental methodology (Section 4.1), and problem formulations (Section 4.2) used to support the motivating analysis presented previously in Section 2 and the more comprehensive analysis presented subsequently in Section 5. The specific algorithms used to solve the sensor placement formulations are described in Section 4.3.

4.1 Networks and Contamination Scenarios

We report computational results for three real, large-scale municipal water distribution networks. The networks are denoted simply as Network1, Network2, and Network3; the identities of the corresponding municipalities are withheld due to security concerns. Network1 consists of roughly 400 junctions, 500 pipes, and a small number of tanks and reservoirs. Network2 consists of roughly 3,000 junctions, 4,000 pipes, and approximately 50 tanks and reservoirs.

Network3 consists of roughly 12,000 junctions, 14,000 pipes, and a handful of reservoirs; there are no tanks or well sources in this municipality. All of the models are skeletonized, although the degree of skeletonization in Network1 and Network2 is much greater than in Network3.

Graphical depictions of Network1, Network2, and Network3 are respectively shown in the upper left, upper right, and lower portion of Figure 3. Each graphic was produced by semi-manually "morphing" or altering (e.g., through pipe lengthening or coordinate translation/rotation) key topological features of the original network structure to further inhibit identification of the source municipalities. Local topologies were largely preserved in this process, such that the graphics faithfully capture the coarse-grained characteristics of the underlying network structures. Sanitized versions of all three networks, in the form of EPANET input files, are freely available from the authors. While these files contain no coordinate information, all data other than that relating to labels (which have been anonymized) are unaltered. Consequently, all computed hydraulic and water quality information accurately reflect (within the fidelity limits of the data and the computational model) the dynamics of the source municipalities. Our goals in making these models available to the broader research community are to facilitate independent replication of our results and to introduce larger, more realistic networks into the currently limited suite of available test problems.

Network hydraulics are simulated over a 96 hour duration, representing four iterations of a typical daily demand cycle. For each junction with non-zero demand, a single contamination event is defined. Each contamination event starts at time $t = 0$ and continues for a duration of 12 hours. Events are modeled as biological mass injections with a constant rate of $5.78e + 10$ organisms per minute. We assume uniform contamination event probabilities, such that all results are normalized by the number of non-zero demand junctions to obtain an expected contamination event impact. Water quality simulations are performed for each event, with a time-step resolution of 5 minutes. The resulting τ_{ej} (as defined in Section 4.2) are then used to compute the impact coefficients d_{ej} for the various design objectives. All hydraulic and water quality simulations are performed using EPANET (Rossman, 2000).

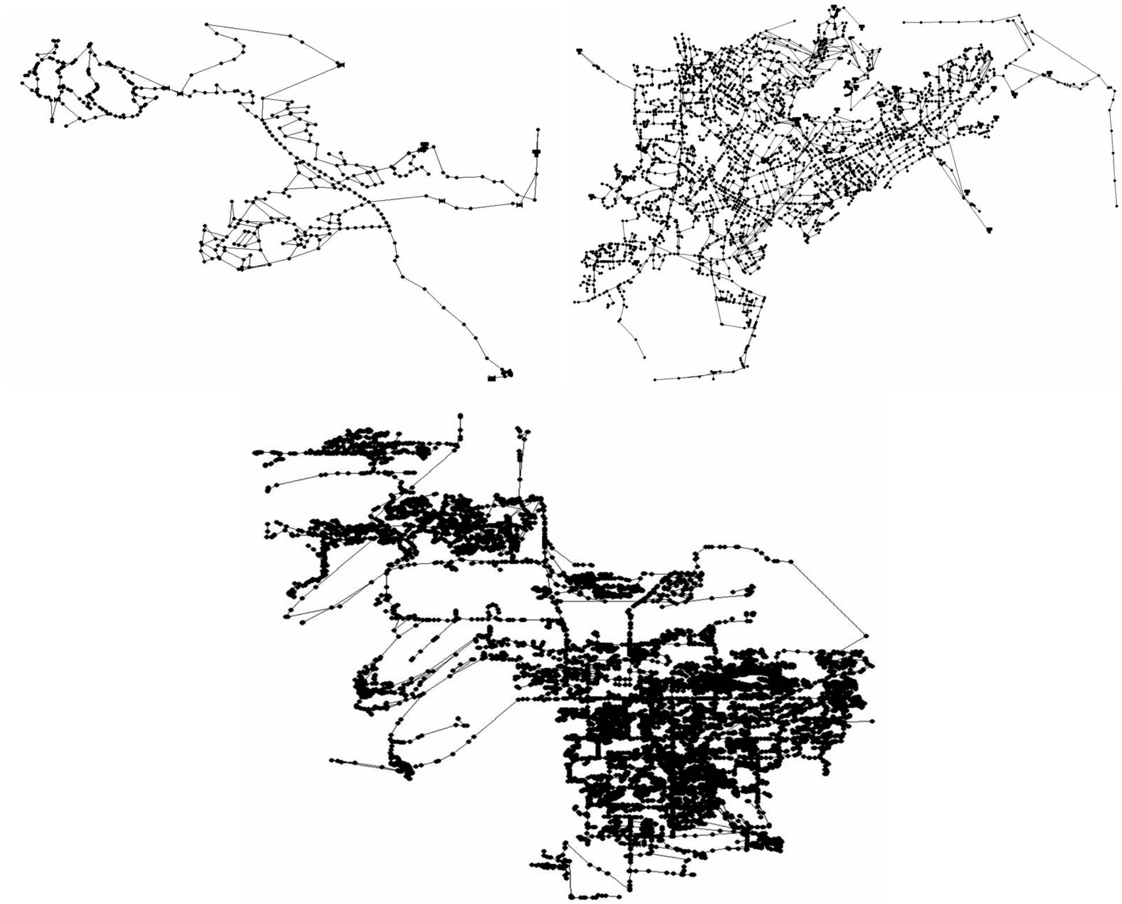

Figure 3: Graphical depictions of Network1 (upper left), Network2 (upper right), and Network3 (lower) municipality distribution topologies.

4.2 Optimization Model

To determine an optimal sensor placement x and the corresponding minimal performance metric $f(x)$, we formulate both the expected-case and robust sensor placement problems as Mixed-Integer (Linear) Programs (MIPs), which we then solve using both problem-specific heuristics and a commercially available MIP solver. The MIP-related terms used throughout this paper are defined in the *Mathematical Programming Glossary* (Greenberg, 2006). As we previously showed in (Berry et al., 2006b), the expected-case sensor placement optimization problem is equivalent to the well-known p-median facility location problem. The MIP formulation of the p-median problem is given as follows, where $\mathcal{E}$ represents the set of contamination events, $\mathcal{L}$ represents the set of network junctions at which a sensor can be placed, p represents the available number of sensors, and q represents a (free) "dummy" sensor that can detect all events given a sufficiently long time horizon (e.g, due to diagnoses at medical facilities):

$$\text{Minimize} \quad \sum_{e \in \mathcal{E}} \sum_{j \in \mathcal{L} \cup \{q\}} d_{ej} x_{ej} \tag{1}$$

$$\text{Subject to} \quad \sum_{j \in \mathcal{L} \cup \{q\}} x_{ej} = 1 \qquad , \forall e \in \mathcal{E} \tag{2}$$

$$x_{ej} \leq y_j \qquad , \forall j \in \mathcal{L}, e \in \mathcal{E} \tag{3}$$

$$\sum_{j \in \mathcal{L}} y_j = p \tag{4}$$

$$y_j \in \{0, 1\} \qquad , \forall j \in \mathcal{L} \tag{5}$$

$$0 \leq x_{ej} \leq 1 \qquad , \forall e \in \mathcal{E}, j \in \mathcal{L} \cup \{q\} \tag{6}$$

The binary y_j variables determine whether a sensor is placed at a junction $j \in \mathcal{L}$. Linearization of the optimization objective is achieved through the introduction of auxiliary variables x_{ej}, which indicate whether a sensor placed at junction j is the first to detect contamination event e. Constraint 3 ensures that detection is possible only if a sensor exists at a junction. The x_{ej} variables are implicitly binary due to a combination of binary y_j, Constraint 3, and the objective function pressure induced by Equation 1. Constraint 4 ensures that exactly p sensors are placed in the network. Constraint 2 guarantees that each contamination event $e \in \mathcal{E}$ is first detected by exactly one sensor, either at q or in the set $\mathcal{L}$; ties are broken arbitrarily. Finally, the objective function (Equation 1) ensures that detection of an event e is assigned to the junction $j \in \mathcal{L} \cup \{q\}$ such that d_{ej} is minimal.

The impact of a potential contamination event is determined via transport simulation. EPANET (Rossman, 2000) is used to generate a time-series τ_{ej} of contaminant concentration at each junction $j \in \mathcal{L}$ for each event $e \in \mathcal{E}$. The resulting time-series are then used to compute the network-wide impact d_{ej} of the event e assuming first detection via a sensor placed at junction j. More formally, let γ_{ej} denote the earliest time t at which a sensor at junction j can detect contamination due to event e, e.g., when contaminant concentration reaches a specific detection threshold. If contaminant from event e fails to reach junction j, then $\gamma_{ej} = t^*$, where t^* denotes either the end of the simulation; otherwise, γ_{ej} can easily be computed from τ_{ej}.

Further, let $d_e(t)$ denote the network-wide damage incurred by an event e up to time t. Next, we define $d_{ej} = d_e(\gamma_{ej})$, i.e., the aggregate, network-wide damage incurred if event e is first detected at time γ_{ej}. In our analysis, $d_{sq} = d_s(t^*)$. We assume without loss of generality that a sensor placed at a junction $j \in \mathcal{L}$ is capable of immediately detecting any contamination from event $e \in \mathcal{E}$ – assuming the contaminant can reach junction j – once non-zero concentration levels of a contaminant are present. In the absence of realistic alarm procedures and mitigation strategies, we assume that both consumption and propagation of contaminant is terminated once detection occurs; extensions to deal with delayed notification are described in (Berry et al., 2006b). Finally, we observe that the p-median optimization formulation – through the use of d_{ej} coefficients – allows for the use of arbitrarily complex contamination scenarios, e.g., multiple simultaneous injection sites with different contaminants at variable injection strengths and durations.

We have also investigated extensions of the basic MIP formulation to robust metrics. While expression of a MIP formulation to minimize *Worst* is a straightforward extension of the expectation-based formulation, the *CVaR* (the continuous approximation to *TCE*, which in general is discretized) formulation is significantly more complicated. For reasons discussed in below in Section 4.3, we do not discuss these formulations herein, and instead refer to Greenberg et al. (2007).

We quantify the impact due to a contamination event as the number of individuals sickened by exposure prior to detection by either a sensor or a sufficient time delay (i.e., detection by the dummy sensor q). The specific computation is defined via the demand-based model (in which contaminant ingestion is proportional to volume of water extracted from a distribution system) described in Murray et al. (2006), and the values for the numerous parameters in the dosage-response computation can be obtained from the authors. The Murray et al. (2006) model yields potentially fractional population counts, but to simplify the presentation we round all reported values to the nearest integral value. Alternative models of population exposure have assumed the availability of population estimates on a per-junction basis (Berry et al., 2005a; Watson et al., 2004). While correcting the obvious deficiency of demand-based models, reliable estimates of population density are generally unavailable.

4.3 Algorithms

We have previously described both heuristic and exact algorithms for solving expectation-based MIP formulations of the sensor placement problem (Berry et al., 2006b). We employed commercially available, state-of-the-art MIP solvers, specifically ILOG's CPLEX 10.0 solver[1], to compute provably optimal solutions. Using various modeling techniques to reduce the size of the basic formulation, we were able to identify optimal solutions to Network3 (our largest test network) in roughly 15 minutes of CPU time on a modern computing workstation. These techniques take advantage of equality in the arrival time of contaminant at various junctions, due to the imposition of a discretized water quality time-step. Consequently, the impacts d_{ej} are identical for various junctions j, which can be collected into "superlocations", thereby reducing the effective size of the formulation (Berry et al., 2007).

We also applied a Greedy Randomized Adaptive Search Procedure (GRASP) to heuristically generate high-quality solutions to the expectation-based MIP formulation. The algorithm, fully described in Resende and Werneck (2004), is a simple multi-start local search procedure in which steepest-descent hill-climbing is applied to a number N of initial solutions. The local search neighborhood used in the GRASP algorithm is based on facility exchange: each move consists of closing a currently opened facility and opening a currently closed facility. The steepest-descent procedure selects the exchange that results in the largest increase in performance at each iteration, and terminates once no improvements are possible. The best of the N solutions is returned by the algorithm. Our experiments indicate that the GRASP heuristic obtains solutions significantly faster than the MIP solves described above, e.g., in under three minutes for Network3. Further, in all cases investigated to date, the obtained solutions were optimal, i.e., equivalent in quality to those obtained by CPLEX.

We extended the GRASP heuristic to enable solution of the robust variants of the MIP formulation described in Section 4.2. The extensions involved modification of the move evaluation code that determines the change in performance associated with simultaneously removing a sensor from junction x and placing it instead at an open junction y. The efficiency of the resulting heuristic is dictated by the speed of move evaluation, which can be accelerated by

[1] http://www.ilog.com

various analytic techniques specific to the p-center and related facility location problems; we defer to Mladenovic et al. (2003) for a discussion of these techniques.

5 Expectation versus Robust Sensor Placements

We now examine the performance differences between expectation-based and robust sensor placements on our test networks. Our analysis is broken into two components. We begin in Section 5.1 by expanding the motivational analysis presented in Section 2 to additional robustness measures and test networks. In Section 5.2, we then discuss several key qualitative differences between expectation-based and robust placements in terms of sensor locations in Network2.

5.1 A Quantitative Analysis of Placement Characteristics

For each of our test networks, we use the heuristic algorithm described in Section 4.3 to develop sensor placements that attempt to independently minimize *Mean* performance and the various robust metrics. As discussed in Section 6, we cannot in general guarantee the optimality of robust sensor placements due to the increased difficulty of the corresponding robust MIP formulations relative to the baseline expectation-based MIP formulation. The performance of each of the resulting sensor placements is then quantified in terms of the *Mean*, *VaR*, *TCE*, and *Worst* metrics. The results for Network1 through Network3 are respectively shown in Tables 1 through 3. We observe that in each of the tables, the inequality *VaR* $\leq$ *TCE* $\leq$ *Worst* holds, as required, for the diagonal entries.

We first consider the results for Network1 (see Table 1), in which 5 sensors are placed to protect against 105 contamination events; contamination events are initiated at each of the 105 out of approximately 400 junctions with non-zero demand. Due to the small scale of this problem, we are able to establish the optimality of the *Worst* sensor placement by exactly solving the MIP formulation; we were unable to establish optimality for the *TCE* sensor placement. Relative to the example shown in Section 2, we observe even more dramatic differences between the *Mean* and *Worst* sensor placements: the worst-case impact can be cut in half for less than a 13% increase in the mean impact. Via exhaustive enumeration of the solution space via a modified

	Performance Metric			
Objective to Minimize	*Mean*	*VaR*	*TCE*	*Worst*
Mean	**143**	476	749	1249
VaR	175	**388**	824	1447
TCE	190	476	**539**	679
Worst	162	565	587	**605**

Table 1: Performance of expectation-based and robust sensor placements in terms of various metrics for Network1, generated using the GRASP heuristic. The placements consist of 5 sensors mitigating against 105 possible contamination events.

MIP branch-and-bound procedure, we determined that there are in fact a number of *alternative global optima* that satisfy $Worst = 605$. This finding raises the possibility that solutions with $Worst = 605$ and $Mean \leq 162$ (13% above the minimal 143 value) may exist. Indeed, using a modified version of our heuristic algorithm that allows for specification of side constraints, we found such a solution with $Worst = 605$ and $Mean = 148$; the latter represents roughly a 3% increase relative to the optimal value of $Mean = 143$. This observation further illustrates the degree to which it is possible to trade off robust versus expectation-based performance; in particular, it seems likely that decision-makers would prefer this particular *Worst* placement over the optimal *Mean* placement.

Although we could in principle perform a similar analysis for each of the results shown in Tables 1 through 3, side constraints further increase the difficulty of the robust MIP formulations, which as discussed in Section 6 is already substantial. Rather, we simply note that optimality (or presumed optimality) with respect to one metric does *not* guarantee conditional optimality (e.g., optimal on a secondary measure given a constraint on a primary measure) on the complementary measures, due to the potential presence of alternative optima. Finally, we observe that although the performance characteristics of the *Mean* and *Worst* placements are significantly different, the placements themselves are not; the two *Worst* placements discussed above locate sensors at respectively two and three of the junctions at which sensors are located in the *Mean* placement.

Given that *VaR*, *TCE*, and *Worst* all quantify related aspects of the distribution of strictly high-impact contamination events, we expected *a priori* that sensor placements minimizing these robustness measures would be strongly correlated in terms of their performance, i.e., sensor placements yielding minimal performance with respect to one robust metric will yield near-

	Performance Metric			
Objective to Minimize	*Mean*	*VaR*	*TCE*	*Worst*
Mean	**685**	2244	2953	4902
VaR	740	**2019**	2699	5076
TCE	757	2112	**2508**	3962
Worst	869	2773	2990	**3490**

Table 2: Performance of expectation-based and robust sensor placements in terms of various metrics for Network2, generated using the GRASP heuristic. The placements consist of 20 sensors mitigating against 1621 possible contamination events.

	Performance Metric			
Objective to Minimize	*Mean*	*VaR*	*TCE*	*Worst*
Mean	**320**	1214	1767	4780
VaR	335	**1188**	1781	5794
TCE	343	1283	**1685**	4219
Worst	463	1934	2315	**3079**

Table 3: Performance of expectation-based and robust sensor placements in terms of various metrics for Network3, generated using the GRASP heuristic. The placements consist of 20 sensors mitigating against 9705 possible contamination events.

minimal performance in terms of all robust metrics. Unexpectedly, the data shown in Table 1 indicate this is not the case. For example, the *Worst* performance of the *VaR*-optimal placement is more than double that of the optimal *Worst* performance. Even discounting potential effects due to alternative global optima, the effect remains significant; minimizing *Worst* subject to *VaR* ≤ 388 yields only a slight reduction in *Worst*, to 1249. Similar discrepancies exist between the observed and optimal values of *TCE* given a *VaR*-optimal placement. Of course, minimization of *VaR* allows for any distribution of the remaining α proportion of high-impact events, so the results are consistent. However, the degree of the divergence was unexpected. In general, this behavior simply reinforces the importance of understanding and analyzing the performance metrics used in optimization; apparently subtle definitional differences (e.g., between *TCE* and *Worst*) in metrics can yield significant differences in both sensor placements and performance.

Next, we consider the results for Network2 (see Table 2), which extends the analysis presented in Section 2 to other robust metrics; we are unable to establish optimality of any of the robust sensor placements for Network2 and Network3. Expanding on the previously noted observation that trade-offs in *Mean* and *Worst* performance are possible, we again observe alternative optima in this problem for the *Worst*-optimal performance. Mirroring the approach discussed above for Network1, we were able to generate a solution via imposition of side constraints with *Worst* $= 3490$ and *Mean* $= 768$, in contrast to the initial value of *Mean* $= 869$ given the *Worst*-optimal solution. Consequently, it is possible in Network2 to obtain a nearly 30% reduction in worst-case impact at the expense of a relatively minor 12% increase in mean impact. Interestingly, despite similar performance, this solution and the *Mean*-optimal solution share sensors at only *two* of the possible twenty junctions in common. Finally, as with the results for Network1, the performance of the robust metrics is not strongly correlated – even accounting for the presence of alternative global optima.

We conclude by noting that results analogous to those observed for Network1 and Network2 extend to Network3, the results for which are shown in Table 3. Overall, our primary conclusions – (1) that it is possible to trade off expectation-based and robust performance and (2) the performance of various robust sensor placements is not strongly correlated – hold over a range of distribution network scales, from very small municipalities to large-scale cities. Consequently, the issues we raise in our analysis are broadly applicable to decision-makers in the water security domain.

5.2 A Qualitative Assessment of Placement Characteristics

Quantitative analysis is only one avenue to understanding and exploring the relationships between expectation-based and robust sensor placements. In this section, we compare and contrast the qualitative characteristics of expectation-based and worst-case sensor placements for Network2, each containing 20 sensors. The locations of the corresponding sensor placements are respectively shown in the left and right sides of Figure 4. In Network2, water is treated at a single source and pumped in stages to successively higher elevations. To compare the two sensor placements, we consider characteristics such as the size and number of pipes connected to the sensor junctions, in addition to the demand at sensor junctions. Further, we consider the number of contamination events that are detected by each sensor placement, the average impact of these contamination events, and the time to detection.

In both placements, the sensors are located at junctions along relatively large diameter pipes, which are additionally often connected to more than two pipes; about half of the sensors are located at junctions with large demand. Specifically, all sensors are located on junctions con-

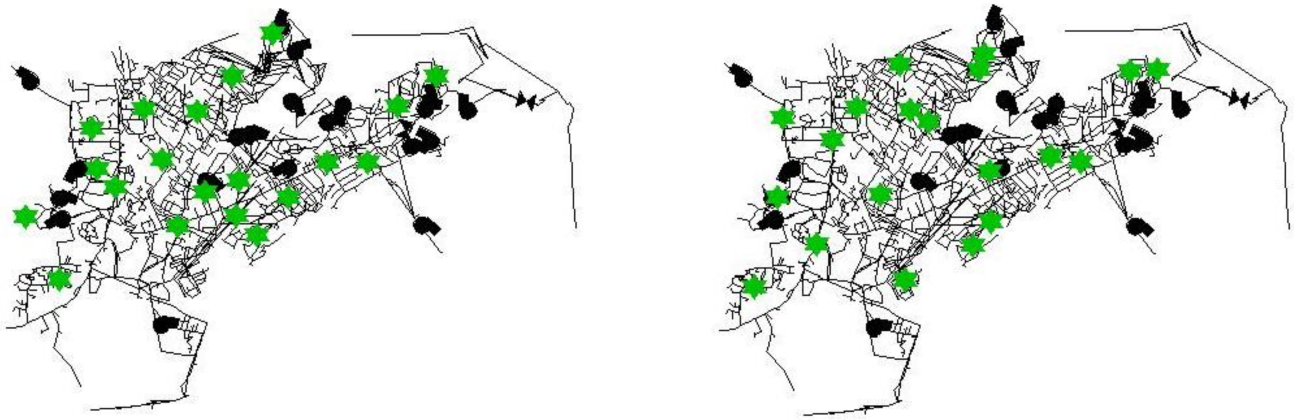

Figure 4: The location of sensors corresponding to *Mean* (left figure) and *Worst* (right figure) sensor placements for Network2. Junctions with sensors are denoted by "star"-shaped graphical overlays.

nected to 8 inch or larger diameter pipes, which is the median diameter of pipes in Network2. Moreover, the majority of sensor-equipped junctions are connected to 12 inch pipes or greater (17-18 of the 20). One difference in the two placements, however, is that the *Worst* placement locates half of the sensors on junctions connected to 20 inch or larger pipes, while only 25% of the sensors in the *Mean* placement are connected to 20 inch pipes or larger.

In both placements, 17 of the 20 sensors are located on junctions connected to 3 or more pipes; none are placed at dead-end junctions. Further, 8 of the 20 sensors in both placements are located on junctions in the top quartile of demands. For the expectation-based placement, 5 sensors are located at zero-demand junctions, while for the *Worst* placement, 8 sensors are located at zero-demand junctions.

It appears from examination of Figure 4 that sensors in the *Worst* placement are somewhat closer together, possibly resulting in less spatial coverage of the distribution network. However, approximately the same number of contamination events are detected by physical sensors: roughly 850 events out of a total of approximately 1,600. Each sensor is responsible for detecting approximately 42 contamination events on average. The average time to detection for each placement is similar; 7 hours for the *Mean* placement and 9 hours for the *Worst* placement. However, the average impact of the contamination events at the time of detection by the *Mean* placement is 685 individuals, in contrast to 882 individuals for the *Worst* placement. It is also interesting to note that a sensor is located much closer to the water source in the *Worst* placement, but not in the *Mean* placement.

In summary, the two sensor placements are surprisingly quite similar in terms of the diameter of connected pipes, the number of connected pipes, the demand at the sensor junctions, and the number of contamination events detected. Notable differences are that the *Worst* solution places more sensors on larger-diameter pipes, does not demonstrate an even spatial spread throughout the network, and has a sensor located much closer to the source. Other salient differences include the observation that the *Worst* placement allows for significantly higher impacts on average (which is obviously necessary to achieve low *Worst* performance), but counterintuitively takes longer on average to detect contamination events. Overall, subtle differences in sensor locations appear to be responsible for the large observed differences in terms of both *Mean* and *Worst* performance.

6 Computational Experience

We now analyze the computational properties of the GRASP heuristic and the MIP models described in Section 4.3, contrasting differences between expected-case and robust optimization models. As hinted at previously, robust MIP formulations are empirically much more difficult to solve than their expectation-based counterparts. To quantify this discrepancy, we consider the average run-times required to generate a single local optimum using the GRASP heuristic for the *Mean*, *VaR*, *TCE*, and *Worst* performance metrics. Our computational platform is a workstation containing 64-bit AMD 2.2GHz Opteron CPUs running the Linux 2.6 operating system; the platform possess 64GB of RAM, such that run-time issues relating to memory paging are non-existent. All codes were written in C++ and compiled using the GNU gcc compiler. The results for all three of our test networks are shown in Table 4, using the sensor budgets indicated in Section 4.1. The run-times include the time required to load the problem instance.

The results clearly illustrate the difficulty of robust variants of the sensor placement problem. Although Network1 run-times are clearly negligible for any metric, the divergence between the *Mean* and other metrics is significant for Network2; the run-times under the *Mean* and *Worst* metrics differ by a factor of 100, and are even larger under the *VaR* and *TCE* metrics. Relative to Network1, the growth in difficulty is accentuated in part due to the growth in the

	Mean Run-Time per Local Optimum		
Objective to Minimize	Network1	Network2	Network3
Mean	0.01s	0.81s	6.5s
Worst	0.02s	97s	4.4hrs
VaR	0.05s	643s	20.4hrs
TCE	0.06s	810s	26.0hrs

Table 4: Mean run-times required for the GRASP heuristic to generate a local optimum to both expectation-based and robust variants of the sensor placement problem, for each of our test networks.

sensor budget p, as the number of exchanges available from any solution is a monotonically increasing function of both $|\mathcal{L}|$ and p for the range of p we consider. Even larger, analogous discrepancies are observed for Network3, where the run-times under the *Mean* and *Worst* metrics differ by a factor of nearly 2,500. The difficulty of computing samples for the *VaR* and *TCE* metrics is much greater than that for *Worst*. This is due to the additional need for sorting the impacts (in the case of *VaR* and *TCE*) and computing the tail expectation (in the case of *TCE*).

We now consider the relative difficulty of expectation-based and robust MIP formulations for exact solvers. Specifically, we executed CPLEX 10.0 on each of our test networks, to independently minimize *Mean*, *CVaR*, and *Worst*. The computational platform was identical to that described above for the heuristic tests, and a limit of 24 (and in some cases greater, for Network3) hours was imposed on each individual run. The results are reported in Table 5.

We first examine the results for Network1, observing that minimization of the robust metrics requires several orders of magnitude more run-time than required for the *Mean* metric. However, minimization of *CVaR* is less costly than *Worst*. We currently have no explanation for this discrepancy. Next, we examine the results for Network2 and Network3. In no case could CPLEX minimize the robust metrics within the allocated time limit. In several cases, a feasible solution could not be located, and in all cases the heuristic solutions yielded better performance than the best solution found by CPLEX. Overall, these results clearly reinforce the dramatic differences in difficulty involved in minimization of expectation-based versus robust performance metrics; the latter require at least 20 times more computational effort, and in most cases, significantly more.

Overall, the data presented in Tables 4 and 5 illustrate the challenges associated with optimization of robust performance metrics. Although MIP methods are tractable in the case of

	Run-Time		
Objective to Minimize	Network1	Network2	Network3
Mean	0.70s	3m2s	47m31s
Worst	8m20s	>24hrs	>48hrs
CVaR	3m18s	>24hrs	>96hrs

Table 5: Run-times to solve the exact MIP models for expectation-based and robust variants of the sensor placement problem, for each of our test networks.

minimizing *Mean* impacts, optimal robust solutions - or at least proofs of optimality - are currently out of reach of exact methods. Even with heuristics, locating high-quality solutions to robust formulations requires a significant computational investment. However, even lacking optimal solutions, the fundamental conclusions presented in Section 5 still hold: it is possible to trade off expected versus robust performance. Future improvements in heuristic and exact technologies will further enhance our ability to exploit this property, and to better understand the relationship between the various robust metrics. Finally, we observe that the relative difficulty of robust optimization is not necessarily inherent. Our results are empirical, rather than theoretical, and it is possible that additional research will expose techniques for significantly improving algorithm performance, e.g., cuts in the case of MIPs or more effective move evaluators in the case of heuristics. Algorithms for minimizing the expected case, i.e., for solving the p-median formulation, have been extensively studied for decades, and only recently have these algorithms yielded results as impressive as those we report.

7 Conclusions

Most extant algorithms for the sensor placement problem in water distribution networks consider minimization of the expected impact of a contamination event. However, the solutions generated by these algorithms admit a number of low-probability, very high-impact contamination events. The presence of these events, in addition to consideration of known inaccuracies in and difficulties associated with contamination event probability estimation, should motivate decision makers to assess the differences between solutions that minimize expected impact and those that focus strictly on high-consequence contamination events. We introduce a number of so-called robust metrics for quantifying the impact of high-consequence contamination events. Using both heuristic and exact optimization algorithms, we then contrast the performance characteristics of solutions that respectively attempt to minimize the mean and robust metrics. We show that it is possible to gain significant reductions in the number and degree of high-consequence events, at the expense of moderate increases in the mean impact of a contamination event. The existence of this trade-off should be of significant interest to CWS designers, given inherent issues involved with event probability estimation and the implicit desire to mitigate against 9/11-style attacks. Additionally, we find that performance with respect to different robust metrics is not highly correlated, further emphasizing the need to develop a deeper understanding of the relationship between solutions developed using different optimization metrics. Finally, we demonstrate that solution of robust formulations of the sensor placement problem are significantly more difficult than for their expectation-based counterpart. Although heuristics can identify high-quality solutions for robust formulations, exact methods are unable to tackle all but the smallest test networks. Non-trivial research effort will ultimately be required to develop truly efficient algorithms for solving, especially to optimality, robust formulations.

Acknowledgments

Sandia is a multipurpose laboratory operated by Sandia Corporation, a Lockheed-Martin Company, for the United States Department of Energy under contract DE-AC04-94AL85000. This work was funded through an Interagency Agreement with the United States Environmental Protection Agency's National Homeland Security Research Center. The authors acknowledge assistance from Harvey Greenberg and Tod Morrison (for model development), in addition to Lee Ann Riesin and Jonathan Berry (for software and algorithmic support).

References

Artzner, P., Delbaen, F., Eber, J., and Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, 9:203–228.

Berry, J., Carr, R., Hart, W., Leung, V., Phillips, C., and Watson, J. (2006a). On the placement of imperfect sensors in municipal water networks. In *Proceedings of the 8th Symposium on Water Distribution Systems Analysis*.

Berry, J., Carr, R., Hart, W., and Phillips, C. (2007). Scalable water network sensor placement via aggregation. In *Proceedings of the ASCE/EWRI Congress*.

Berry, J., Fleischer, L., Hart, W., Phillips, C., and Watson, J. (2005a). Sensor placement in municipal water networks. *Journal of Water Resources Planning and Management*, 131(3):237–243.

Berry, J., Hart, W., Phillips, C., Uber, J., and Walski, T. (2005b). Water quality sensor placement in water networks with budget constraints. In *Proceedings of the ASCE/EWRI Congress*.

Berry, J., Hart, W., Phillips, C., Uber, J., and Watson, J. (2006b). Sensor placement in municipal water networks with temporal integer programming models. *Journal of Water Resources Planning and Management: Special Issue on Drinking Water Distribution Systems Security*, 132:218–224.

Carr, R., Greenberg, H., Hart, W., Konjevod, G., Lauer, E., Lin, H., Morrison, T., and Phillips, C. (2006). Robust optimization of contaminant sensor placement for community water systems. *Mathematical Programming*, 107(1):337–356.

Greenberg, H. (1996–2006). *Mathematical Programming Glossary*. World Wide Web, `http://www.cudenver.edu/~hgreenbe/glossary/`.

Greenberg, H., Morrisson, T., and Hart, W. (2007). Robust MIP formulations for sensor placement optimization. Technical report, Sandia National Laboratories.

Holton, G. A. (2003). *Value-at-Risk: Theory and Practice*. Academic Press.

Kessler, A., Ostfeld, A., and Sinai, G. (1998). Detecting accidental contaminations in municipal water networks. *Journal of Water Resources Planning and Management*, 124(4):192–198.

Mladenovic, N., Labbe, M., and Hansen, P. (2003). Solving the p-center problem with tabu search and variable neighborhood search. *Networks*, 42(1):48–64.

Murray, R., Uber, J., and Janke, R. (2006). Modeling acute health impacts resulting from ingestion of contaminated drinking water. *Journal of Water Resources Planning and Management: Special Issue on Drinking Water Distribution Systems Security*, 132:293–300.

Ostfeld, A. and Salomons, E. (2004). Optimal layout of early warning detection stations for water distribution systems security. *Journal of Water Resources Planning and Management*, 130(5):377–385.

Resende, M. and Werneck, R. (2004). A hybrid heuristic for the p-median problem. *Journal of Heuristics*, 10(1):59–88.

Rossman, L. (2000). The EPANET programmer's toolkit for analysis of water distribution systems. In *Proc. of the Annual Water Resources Planning and Management Conference*.

USEPA (2005). WaterSentinel System Architecture. Technical report, U.S. Environmental Protection Agency.

Watson, J., Greenberg, H., and Hart, W. (2004). A multiple-objective analysis of sensor placement optimization in water networks. In *Proceedings of the ASCE/EWRI Congress*.